

W 1579

## PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-198528

(43)Date of publication of application : 31.07.1998

(51)Int.Cl.

G06F 3/06

G06F 3/06

G06F 3/06

G06F 12/16

(21)Application number : 09-001274

(71)Applicant : HITACHI LTD

(22)Date of filing : 08.01.1997

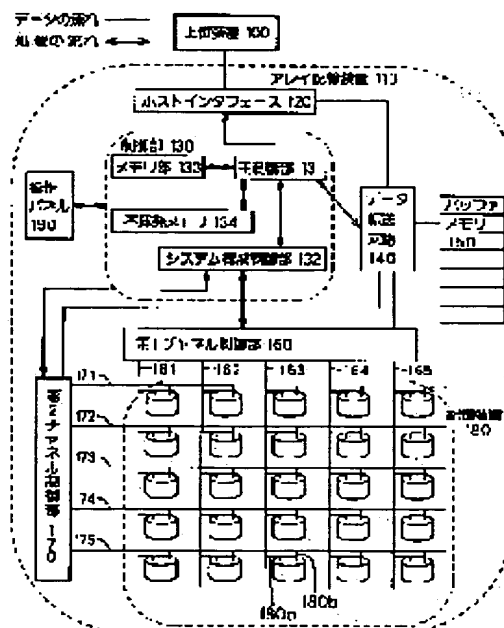
(72)Inventor : IKEDA MASANOBU  
KOBAYASHI MASAOKI  
TAKAMOTO KENICHI

## (54) ARRAY STORAGE DEVICE AND ITS CONTROLLING METHOD

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To improve fault resistance performance and data input and output performance in an array storage device, and to attain the simple change of a system constitution.

**SOLUTION:** Each of plural storage devices 180 arranged like a matrix is provided with plural interface ports 180a and 180b, and connected through column directional buses 161-165 and line directional buses 171-175 with a first channel control part 160 and a second channel control part 170. Then, the plural storage devices 180 constituting the parity group of an RAID are selected so as to be exclusively connected with each of the buses 161-165 and 171-175 so that parallel access to the plural storage devices 180 in the parity group and the multiplexing of the bus connection paths can be realized. The selected result of the storage device 180 is stored in a non-volatile memory 134 of a control part 130, and the initialization/change of the parity group is operated through a control panel 190 as necessary.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the  
examiner's decision of rejection or application  
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of  
rejection][Date of requesting appeal against examiner's decision  
of rejection]

[Date of extinction of right]



(19) 日本国特許庁 (J P)

(12) 公開特許公報 (A)

(11) 特許出願公開番号

特開平10-198528

(43) 公開日 平成10年(1998) 7月31日

(51) Int.Cl. <sup>8</sup>	識別記号	F I	
G 0 6 F 3/06	5 4 0	G 0 6 F 3/06	5 4 0
	3 0 2		3 0 2 J
	3 0 4		3 0 4 C
12/16	3 2 0	12/16	3 2 0 L

審査請求 未請求 請求項の数 3 O L (全 12 頁)

(21) 出願番号 特願平9-1274

(22) 出願日 平成9年(1997) 1月8日

(71) 出願人 000005108

株式会社日立製作所

東京都千代田区神田駿河台四丁目6番地

(72) 発明者 池田 雅信

神奈川県小田原市国府津2880番地 株式会

社日立製作所ストレージシステム事業部内

(72) 発明者 小林 正明

神奈川県小田原市国府津2880番地 株式会

社日立製作所ストレージシステム事業部内

(72) 発明者 ▲高▼本 賢一

神奈川県小田原市国府津2880番地 株式会

社日立製作所ストレージシステム事業部内

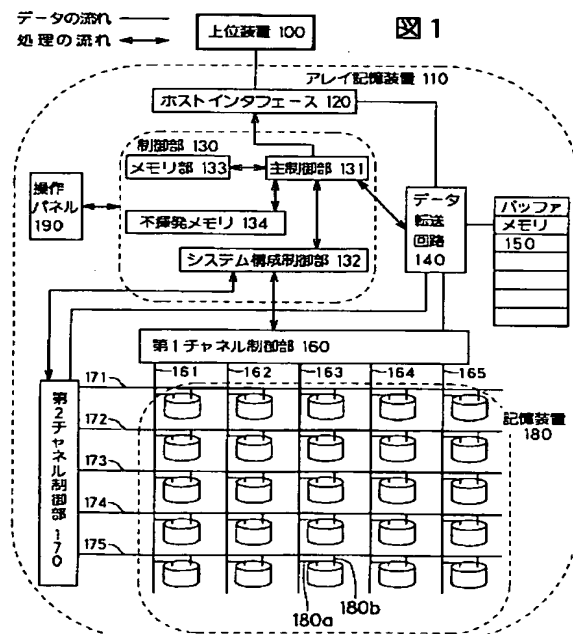
(74) 代理人 弁理士 筒井 大和

(54) 【発明の名称】 アレイ記憶装置およびその制御方法

(57) 【要約】

【課題】 アレイ記憶装置における耐故障性能およびデータ入出力性能の向上と、システム構成の容易な変更を可能にする。

【解決手段】 マトリックス状に配置される複数の記憶装置180の各々は、複数のインタフェースポート180a、180bを持ち、列方向のバス161~165および行方向のバス171~175を介して第1チャンネル制御部160および第2チャンネル制御部170に接続され、バス161~165およびバス171~175の各々に対して互いに排他的接続されるように、RAIDのバリティグループを構成する複数の記憶装置180を選択することでバリティグループ内の複数の記憶装置180に対する並列的なアクセスと、バス接続経路の多重化を実現し、記憶装置180の選択結果は、制御部130の不揮発メモリ134に格納され、操作パネル190を介して随意にバリティグループの初期設定/変更が行われる。



## 【特許請求の範囲】

【請求項1】 各々が複数の入出力ポートを備えた複数の記憶装置と、

少なくとも一つの前記記憶装置の複数の前記入出力ポートの各々に接続される複数のバスと、

複数の前記バスが、当該バスが接続される前記記憶装置の前記入出力ポート毎に接続され、接続された前記バスを介して前記記憶装置に対して並列的な入出力処理を実行する複数の入出力制御部と、

複数のデータブロックおよび当該データブロックから生成される冗長データブロックが分散して格納され、論理的なパリティグループを構成する複数の前記記憶装置の個々の前記入出力ポートが、可能な限り、複数の前記バスに対して互いに排他的に接続されるように前記パリティグループを複数の前記記憶装置に論理的に割り当てるシステム構成制御手段と、

を備えたことを特徴とするアレイ記憶装置。

【請求項2】 請求項1記載のアレイ記憶装置において、

前記システム構成制御手段は、前記記憶装置の数量の増減および前記パリティグループの設定変更要求の少なくとも一方に応じて、前記パリティグループに属する前記記憶装置の組み合わせを論理的に任意に変更する機能を備えたことを特徴とするアレイ記憶装置。

【請求項3】 各々が複数の入出力ポートを備えた複数の記憶装置と、

少なくとも一つの前記記憶装置の複数の前記入出力ポートの各々に接続される複数のバスと、

複数の前記バスが、当該バスが接続される前記記憶装置の前記入出力ポート毎に接続され、接続された前記バスを介して前記記憶装置に対して並列的な入出力処理を実行する複数の入出力制御部と、

を含むアレイ装置の制御方法であって、

複数のデータブロックおよび当該データブロックから生成される冗長データブロックが分散して格納される複数の前記記憶装置によって論理的なパリティグループを構成するとき、前記パリティグループに含まれる個々の前記記憶装置の個々の前記入出力ポートが、可能な限り、複数の前記バスに対して互いに排他的に接続されるように前記パリティグループを構成する複数の前記記憶装置を選択することを特徴とするアレイ記憶装置の制御方法。

## 【発明の詳細な説明】

## 【0001】

【発明の属する技術分野】本発明は、アレイ記憶装置およびその制御技術に関し、特に、情報処理システムにおけるデータ記憶装置として使用され、高い信頼性と入出力要求に対する高速な応答が要求されるアレイ記憶装置等に適用して有効な技術に関する。

## 【0002】

【従来の技術】アレイ記憶装置に対する入出力要求の高速応答を可能とするためにバッファメモリを設け、上位装置からの入出力要求をなるべく高速なメモリへのアクセスにて行うことが一般的である。データ読み取り要求に対してバッファメモリ上にアクセス対象のデータが存在しない場合、データ書き込み要求発生時や記憶装置上のデータとバッファメモリ上のデータとの同期を取る場合にバッファメモリと記憶装置との間でデータの送受信が実行される。

【0003】記憶装置の配置／接続方式については従来のアレイ記憶装置では、個々の記憶装置が（１）シングルポートを持って複数の単一バスによる列／行の構成を取るか、（２）複数のインタフェースポートを持って複数のチャンネルによる列／行の構成を取っていた。（１）の構成においては高速性と高信頼性の両立に対して考慮されておらず、（２）の方式ではパリティグループ内の記憶装置の割り当てについて考慮されていない。また、（１）、（２）どちらについても、パリティグループ内の記憶装置数を可変とする制御について、バス障害に対する高い信頼性を持つことについて十分な考慮がされていない。

【0004】特開平6-161672号公報には様々な個数の記憶装置に対するデータ分配の方式について述べられている。この方法では記憶装置の構成変化に対して、ソフトウェアによる設定で対応することが可能である。しかし、この技術において各記憶装置が複数のインタフェースポートを持つ場合について、例えばバス障害に対するアレイ記憶装置の信頼性については考慮されていない。また、複数の列／行にわたる複数の記憶装置への同じアクセスを効率よく行う方式についても考慮されていない。さらに、複数の行／列にまたがる複数の記憶装置を同一パリティグループとして制御する方式についても考慮されていない。

【0005】一方、特開平5-181611号公報の技術では直交するチャンネル接続を持つ記憶装置を配置する構成をとっているが、2通りのアクセス手段による分散アクセスが可能であることとチャンネル障害に対して障害チャンネルを回避して残ったチャンネルを用いてアクセスすることなど、直交チャンネル接続構成でなくても実現できる特徴について言及しているに留まっている。たとえば、記憶装置配置形式についての最適化や、パリティグループ内の複数の記憶装置への同時アクセスを実現する手段について考慮されていない。

## 【0006】

【発明が解決しようとする課題】最近の市場動向として、高速かつ高い信頼性を持つディスクアレイ装置を接続した多目的ファイルサーバに対する需要が高まってきている。

【0007】クライアントに対して複数のディスクイメージを見せることができることから、安定した高速性を

要求されるマルチメディアデータアクセス処理と、高い信頼性を伴った重要データに対するトランザクション処理への対応を要求される。そのためには、アレイ記憶装置内部の記憶装置に対して、高い信頼性と高速性能の実現が必須である。

【0008】従来の技術では、バス障害に対して高い信頼性を保証しつつ、パリティグループ内の記憶装置に対する同時アクセスを実現する方式について考慮されていなかった。

【0009】本発明の目的は、高い信頼性とデータアクセス性能が要求されるアレイ記憶装置において、記憶装置を接続するバスの障害に対する高い信頼性を保持しつつ、パリティグループへのアクセス性能を向上させることにある。

【0010】本発明の他の目的は、高い耐故障性能およびアクセス性能を維持しつつ、パリティグループや装置構成の柔軟な変更を容易に実現することにある。

【0011】

【課題を解決するための手段】本発明では、複数のインタフェースポートを持つ記憶装置を採用し、パリティグループを構成する複数の記憶装置の各々を可能な限り排他的に異なる組み合わせのバスに接続するシステム構成を採用する。そして、データ領域（ブロック）とそこから生成されるパリティ領域（ブロック）の集合であるパリティグループを任意の記憶装置の集まりに論理的に割り当てる手段を設け、記憶装置やパリティグループ等の変更や追加時には、記憶装置に対するパリティグループの割り当ての再組み替えを行う手段を設ける。また、上記各手段についての設定や参照を装置付属の操作パネルから設定する手段を設ける。

【0012】上述のような構成の本発明のアレイ記憶装置では、たとえば、使用者により操作パネルからRAIDレベル、およびパリティグループ内の記憶装置数を随意に設定させ、これらの設定に基づいて、パリティグループを構成する複数の記憶装置の各々が、可能な限り、排他的に異なる組み合わせのバスに接続されるように、パリティグループを記憶装置の集まりに自動的に割り当てる手段により、各々のパリティグループ内の記憶装置に接続されるバスに競合が起きず、記憶装置へ同時にアクセスできるようなアレイ記憶装置の内部構成を自動的に設定できるようにする。また、使用者により操作パネルからパリティグループの増設が指示されたときは、パリティグループの割り当ての再組み替えを行う手段の起動により、記憶装置とパリティグループとの対応付けが再構成され、上記同様のアレイ記憶装置の内部構成が自動的に設定される。

【0013】

【発明の実施の形態】以下、本発明の実施の形態を図面を参照しながら詳細に説明する。

【0014】図1は、本発明の実施の形態のアレイ記憶

装置およびその制御方法が適用される情報処理システムの全体構成の一例を示した概念図である。

【0015】本実施の形態の情報処理システムは、アレイ記憶装置110を上位装置100に接続した構成となっている。

【0016】アレイ記憶装置110は、ホストインタフェース120、制御部130、データ転送回路140、バッファメモリ150、第1チャンネル制御部160、第2チャンネル制御部170、マトリックス状に配列される複数の記憶装置180、操作パネル190より構成される。制御部130は通常マイクロプロセッサと制御用ソフトウェアにより実現され装置全体を制御する主制御部131、システム構成制御部132からなる。制御部130上のメモリ部133には主に入出力要求に関する各管理テーブルを置く。また制御部130は不揮発メモリ134を持つ。不揮発メモリ134には各種動作パラメータを置く。

【0017】記憶装置180として、たとえば、2つのインタフェースポート180aおよびインタフェースポート180bを持つ小型磁気ディスク装置を使用する。小型磁気ディスク装置はその各々のインタフェースポート180aおよび180bを異なる第1チャンネル制御部160および第2チャンネル制御部170に接続しRAID構成を組む。RAIDを構成する記憶装置180の各列はバス161～165の各々にインタフェースポート180aを共通に接続されることによって第1チャンネル制御部160との間におけるデータ入出力が並行的に行われ、各行はバス171～175にインタフェースポート180bを共通に接続されることによって第2チャンネル制御部170との間におけるデータ入出力が並行的に行われる。

【0018】アレイ記憶装置110内はRAIDを構成する記憶装置180の数が異なる1つまたは複数のパリティグループからなる。上位装置100からのアクセス要求は様々であり、ランダムデータアクセス、シーケンシャルデータアクセス、大容量アクセス、小容量データアクセスと多岐にわたる。一般にアレイ記憶装置110は上位装置100からのデータ読み込み要求に対して当該データがバッファメモリ150に存在しない場合、また書き込み要求の後に記憶装置180とのデータの同期を取る場合において、バッファメモリ150と記憶装置180との間でデータの送受信処理が実行される。

【0019】バッファメモリ150から上位装置100がデータをアクセスする処理については上記の記憶装置180へのアクセスと並行して行う。

【0020】図2は記憶装置の接続個数に基づいて記憶装置を配置する方式の一例について示したものである。第1チャンネル制御部160および第2チャンネル制御部170からの直接のアクセスの効率を高めるために、記憶装置配置方法200で示すように、接続位置を記憶装置

接続領域201～記憶装置接続領域209にわけて、記憶装置接続領域201から順に209まで記憶装置180を配置していく。

【0021】すなわち、本実施の形態の記憶装置配置方法200では、各記憶装置180は、互いに列および行方向にて重複することなく排他的にバスに接続されることとなり、各記憶装置180への第1チャンネル制御部160および第2チャンネル制御部170から入出力の実行時にバスの競合は発生せず、しかも、データ入出力経路となるバスは列方向および行方向にて二重に確保される。

【0022】アレイ記憶装置110に実装可能な記憶装置180の最大数より少ない記憶装置180を実装する場合も、記憶装置配置方法200に基づいて、記憶装置180を配置する。その他、実装順の例として、201→202→209→204→207→206→205→208→203などがある。

【0023】ここで、図3に、記憶装置の配置による接続バス数の関係について4つの例を示す。記憶装置配置方法の評価210に例示された配置方法では、記憶装置211、記憶装置212、記憶装置213に接続されるバスの合計は6になる。記憶装置配置方法の評価214に例示された配置方法では、記憶装置215、記憶装置216、記憶装置217に接続されるバスの合計は5になる。記憶装置配置方法の評価218に例示された配置方法では、記憶装置219、記憶装置220、記憶装置221に接続されるバスの合計は4になる。記憶装置配置方法の評価222に例示された配置方法では、記憶装置223、記憶装置224、記憶装置225に接続されるバスの合計は4になる。すなわち、接続する記憶装置配置方法の評価210のように記憶装置を配置することにより、アクセス可能なバスを最大にすることができる。

【0024】また、図4に例示されるように、記憶装置接続例226のようにアレイ記憶装置110に実装する記憶装置数が8のときは、記憶装置接続領域201に記憶装置を5台、記憶装置接続領域202に3台の記憶装置を割り当てることにより、最大で第1チャンネル制御部160のバス161～バス165、第2チャンネル制御部170のバス171～バス175の合計10のバスから、各々同時にアクセスすることが可能となる。

【0025】次に、データ領域とパリティ領域のまとまりであるパリティグループへの記憶装置180の割り当てについて説明する。図3に例示された記憶装置180の配置の場合と同様に、パリティグループに割り当てられる記憶装置180を適当に選ぶことで、パリティグループへ同時にアクセスできる接続バスの数を最大とすることができる。例えば、図5、図6および図7にて、3D+1P（RAID定義例301）（データ（D）領域：パリティ（P）領域＝3：1（グループ内領域割

り当て例302））となるパリティグループを2つ設定した場合について、記憶装置の配置方式と接続バスの関係について3つの例を示す。

【0026】図5に例示されたパリティグループ割り当て例300では、パリティグループ303に記憶装置305～記憶装置308を割り当て、パリティグループ304に記憶装置309～記憶装置312を割り当てる。

【0027】図6に例示されたパリティグループ割り当て例313では、パリティグループ314に記憶装置316～記憶装置319を割り当て、パリティグループ315に記憶装置320～記憶装置323を割り当てる。

【0028】図7に例示されたパリティグループ割り当て例324では、パリティグループ325に記憶装置327～記憶装置330を割り当て、パリティグループ326に記憶装置331～記憶装置334を割り当てる。

【0029】それぞれの独立な接続バス数は、図8のパリティグループ別接続バス比較例335に示すようにパリティグループ325とパリティグループ326の8が最大となる。また、RAIDレベルが同一であり、かつパリティグループ内の記憶装置数が同一である複数のパリティグループをまとめて一つの論理デバイスとして上位装置100からアレイ記憶装置110にアクセスする場合に、上位装置100からの1つの入出力要求が複数のパリティグループに分割されて処理される場合がある。例えば図7のパリティグループ割り当て例324において、上位装置100からパリティグループ325とパリティグループ326を一つの論理デバイスとして認識しているとき、記憶装置327～記憶装置334までの8台にアクセス要求が同時に発生する場合がある。論理デバイスとして認識されているときの同時アクセス可能バス数について、パリティグループ割り当て例300、パリティグループ割り当て例313、パリティグループ割り当て例324のそれぞれの値を、パリティグループ別接続バス比較例335に示す。パリティグループ別接続バス比較例335によれば、パリティグループ割り当て例324の、論理デバイスとして認識されているときの同時アクセス可能バス数が最大の10となっている。したがって、パリティグループ別接続バス比較例335が、各々の記憶装置が持つインタフェースポートをもっとも多く同時に使用することができる記憶装置180の配置例であることがいえる。

【0030】図9および図10を用いてパリティグループの初期設定方法の一例を説明する。まず操作パネル190にてパリティグループを設定するために、各々のパリティグループ501（0，1，2）対応にRAIDレベル502とグループ内記憶装置数503を入力し（ステップ401）、不揮発メモリ134のパリティグループ初期設定テーブル500に登録する（ステップ402）。設定終了か否かを判定し、他のパリティグループを設定する場合は、開始点（ステップ400）へ、そう

でない場合は終了点(ステップ404)へ処理を移す(ステップ403)。パリティグループ初期設定テーブル500は、パリティグループ(501)毎に持ち、項目として、例えば図10に示すようにRAIDレベル502、当該パリティグループのグループ内記憶装置数503を持つ。

【0031】図11および図12にて、パリティグループ初期設定テーブル500を入力として、パリティグループを構成する記憶装置180の配置の決定方法の一例を示す。パリティグループ初期設定テーブル500を入力として(ステップ600)、各パリティグループを構成する記憶装置数Xを調べ、Xの降順(Xの大きいものから順)に記憶装置180にパリティグループ内記憶装置を割り当てていく(ステップ601)。

【0032】未割り当て記憶装置の中から1台目の記憶装置(M1とする)を当該パリティグループに割り当てる(ステップ602)。

【0033】以下、i台目の記憶装置Miの割り当て(ステップ603)について説明する。

【0034】割り当て処理(ステップ604)にて、Mjの第1チャンネル制御部160の接続バスをXj、第2チャンネル制御部170の接続バスをYjとし、Mj(Xj, Yj)と定義し(ステップ605)、未割り当て記憶装置の中から、すでにそのパリティグループに割り当てられたすべての記憶装置Mjと比較して、第1チャンネル制御部160および、第2チャンネル制御部170への2本の接続バスが異なる記憶装置を選び、Miとする。すなわち、未割り当て記憶装置の中から、Xi≠XjかつYi≠Yjとなる記憶装置Mi(Xi, Yi)を当該パリティグループに割り当てる(ステップ606)。割り当て処理(ステップ604)でパリティグループ内記憶装置をすべて割り当てられなかったときには、割り当て処理(ステップ607)で記憶装置の割り当てを行う。未割り当て記憶装置の中から、接続バス競合情報テーブル613の割り当て記憶装置数614がもっとも少ない組み合わせの接続バスに接続された記憶装置Mi(Xi, Yi)を接続バスの昇順に当該パリティグループに割り当てる(ステップ608)。

【0035】以上の処理を当該パリティグループ割り当て記憶装置数分繰り返し(ステップ609)、当該パリティグループ内記憶装置割り当てが終了したとき(ステップ610)、設定を要する別パリティグループがあれば、1台目の記憶装置M1の割り当て処理(ステップ602)から繰り返す(ステップ611)。すべてのパリティグループの設定が終了したときに、図18に示した例のようなパリティグループ割り当てテーブル700を作成して記憶装置番号701を設定し、不揮発メモリ134に登録し、パリティグループを構成する記憶装置180の配置の決定処理を終了する(ステップ612)。

【0036】ここで、前述の割り当て処理(ステップ6

04)の詳細について説明する。まず、接続バス競合情報の管理方式の一例として、パリティグループ設定毎に、図13に例示される接続バス競合情報テーブル613を作成する。例えば、Mi(162, 172)を割り当てたとき、バス162とバス172の割り当て記憶装置数614を0から1へカウントアップする(操作615)。割り当て処理(ステップ604)では、割り当て記憶装置数614が0であるような接続バスの組み合わせを持つ記憶装置Mi(Xi, Yi)を接続バスの昇順に選び、当該パリティグループに割り当てる。

【0037】次に、割り当て処理(ステップ604)でパリティグループ内記憶装置をすべて割り当てられなかったときの処理である割り当て処理(ステップ607)の詳細について説明する。例えば、図14に例示される割り当て例616のように記憶装置M(161, 171)、記憶装置M(162, 172)、記憶装置M(163, 173)、記憶装置M(164, 174)が当該パリティグループにすでに割り当てられているとする。このとき、接続バス競合情報は、図15に例示される接続バス競合情報テーブル617のようになる。未割り当て記憶装置の中から、接続バス競合情報テーブル617の割り当て記憶装置数618がもっとも少ない組み合わせの接続バスに接続された記憶装置Mi(Xi, Yi)を接続バスの昇順に当該パリティグループに割り当てる。すなわち、次の記憶装置割り当て候補はM(161, 172)となる。また、このときの接続バス競合情報は、図16に例示される接続バス競合情報テーブル619のようになる。さらに、次の記憶装置割り当て候補は、割り当て記憶装置数620がもっとも少ない組み合わせの接続バスに接続された記憶装置を昇順に選択して、M(162, 173)となる。このとき、図17に例示されるように接続バス競合情報テーブル621で、バス162、バス173の割り当て記憶装置数622はそれぞれ1から2へカウントアップされる(操作623)。

【0038】上述のような本実施の形態の方法でパリティグループに記憶装置180を割り当てることにより、パリティグループ内に属する複数の記憶装置180が、当該パリティグループ内の他の記憶装置180とは異なるバス161~165、バス171~175に可能な限り排他的に接続されるようにすることで、第1チャンネル制御部160、第2チャンネル制御部170を効率よく使用し、パリティグループへのアクセス要求に対してなるべく多くの記憶装置180への並列的な入出力処理を実現することが可能となる。

【0039】次に、パリティグループの設定が終了し、アレイ記憶装置110が動作しているときの記憶装置180への実際のアクセス経路の確保について説明する。このアクセス経路の確保は、例えば、図19のチャンネル使用状況管理テーブル800と図20のチャンネル記憶装置対応テーブル900にて行う。パリティグループに対してアクセス要求が生じた場合は当該記憶装置180

について、チャンネル記憶装置対応テーブル900にて接続バス/チャンネル番号を調べる。その後、チャンネル使用状況管理テーブル800にてチャンネル/バスの使用状況を調べ、複数の空いているバスを選択してアクセスを行う。このとき、空きバスがないときは資源待ちとなるが、同じパリティグループ内でのバス確保競合の発生を抑えることができる。

【0040】また、特定の記憶装置180の障害が発生し、接続した両方のバスが使用できなくなった場合は、主制御部131でその記憶装置に接続されているバスの番号を判定し、当該バスからのアクセスを抑止するためのフラグを立てる。たとえば、図21に示した例のように、特定の記憶装置1001に障害が発生した場合、第1チャンネル制御部160からバス162を使用不可とし、第2チャンネル制御部170からバス172を使用不可とし、障害の記憶装置1001を迂回するように制御し他の記憶装置へのアクセス経路を確保することができる。例えば記憶装置1002に対しては第1チャンネル制御部160からバス163が使用可能であり、記憶装置1003に対しては第2チャンネル制御部170からバス173が使用可能となる。さらに、障害の記憶装置1001を含んだパリティグループは縮退状態となるものの閉塞状態にはならないので、パリティグループへのアクセスは可能であり、システムの継続使用が可能となる。上記の処理指示は本アレイ記憶装置の使用者により操作パネル190から操作可能である。

【0041】次に、本実施の形態において、すでに定義済みのパリティグループの変更および追加等を実施する場合の一例について説明する。あらたにパリティ用記憶装置とデータ用記憶装置のまとまりであるパリティグループを追加するときには、図10に一例を示したパリティグループ初期設定テーブル500にあらたに追加するパリティグループを設定し、図11および図12に示した記憶装置180へのパリティグループの割り当て処理を起動する。そして、パリティグループ初期設定テーブル500と、パリティグループ割り当てテーブル700を不揮発メモリ134に登録する。ここで、パリティグループ割り当てテーブル700に基づいて記憶装置180へのパリティグループの再割当を実行するために、記憶装置180間で適宜にデータコピーを実行し、既に存在していたパリティグループのデータを保証する。上記の処理指示はユーザにより操作パネル190から操作可能である。

【0042】以上説明したように、本実施の形態のアレイ記憶装置およびその制御方法によれば、マトリックス状に配列された複数の記憶装置に、複数のデータ領域および当該データ領域から生成された冗長データ領域を複数の記憶装置に分散して格納するRAIDを実現する場合において、任意のパリティグループに含まれ、二つのインタフェースポート180aおよび180bを備えた

記憶装置180が、列方向のバス161~165および行方向のバス171~175に対して可能な限り互いに排他的に接続されるようにしたので、任意のパリティグループへの入出力要求発生時に、当該パリティグループを構成する複数の記憶装置180に同時に並列的なアクセスが可能になり、処理の高速化を実現することができる。

【0043】また、データ転送回路140と個々の記憶装置180との間におけるデータ転送経路が、複数の第1チャンネル制御部160および第2チャンネル制御部170、さらには列方向の複数のバス161~165、行方向のバス171~175にて多重に確保されているので、第1チャンネル制御部160、第2チャンネル制御部170、バス161~165、バス171~175等における障害の発生等に際して、各記憶装置180にアクセス不能になる確率が減少し、データ転送経路等における障害発生時のシステムの信頼性を高めることができる。

【0044】また、パリティグループ割り当てテーブル700等を不揮発メモリ134に設定し、操作パネル190等を介して当該パリティグループ割り当てテーブル700等を変更することにより、記憶装置180の台数や、パリティグループ数の増減等に応じて、ユーザが、容易かつ柔軟にパリティグループの割り当て状態を変更することができ、RAIDシステムの多様な管理および運用を実現することができる。

【0045】以上本発明者によってなされた発明を実施の形態に基づき具体的に説明したが、本発明は前記実施の形態に限定されるものではなく、その要旨を逸脱しない範囲で種々変更可能であることはいうまでもない。

【0046】たとえば、上述の実施の形態の説明では、一例として、2次元のマトリックス状に論理的に複数の記憶装置を配置した構成について説明したが、各々がN個(Nは3以上の自然数)の入出力ポートを備えた複数の記憶装置を、N次元以上のマトリックス状に論理的に配置した場合において、各パリティグループ内の複数の記憶装置の各々の各入出力ポートが、可能な限り、互いに排他的に異なるバスに接続される構成とすることも本発明に含まれる。

【0047】

【発明の効果】本発明のアレイ記憶装置によれば、高い信頼性とデータアクセス性能が要求されるアレイ記憶装置において、記憶装置を接続するバスの障害に対する高い信頼性を保持しつつ、パリティグループへのアクセス性能を向上させることができる、という効果が得られる。

【0048】また、本発明のアレイ記憶装置によれば、高い耐故障性能およびアクセス性能を維持しつつ、パリティグループや装置構成の柔軟な変更を容易に実現することができる、という効果が得られる。

【0049】また、本発明のアレイ記憶装置の制御方法



によれば、高い信頼性とデータアクセス性能が要求されるアレイ記憶装置において、記憶装置を接続するバスの障害に対する高い信頼性を保持しつつ、パリティグループへのアクセス性能を向上させることができる、という効果が得られる。

【0050】また、本発明のアレイ記憶装置の制御方法によれば、高い耐故障性能およびアクセス性能を維持しつつ、パリティグループや装置構成の柔軟な変更を容易に実現することができる、という効果が得られる。

【図面の簡単な説明】

【図1】本発明の一実施の形態のアレイ記憶装置およびその制御方法が適用される情報処理システムの全体構成の一例を示した概念図である。

【図2】本発明の一実施の形態のアレイ記憶装置およびその制御方法において、記憶装置の接続個数に基づいて記憶装置を配置する方式の一例を示す概念図である。

【図3】本発明の一実施の形態のアレイ記憶装置およびその制御方法における記憶装置の配置方法の評価の一例を示す概念図である。

【図4】本発明の一実施の形態のアレイ記憶装置およびその制御方法における記憶装置接続方法の一例を示す概念図である。

【図5】アレイ装置における複数の記憶装置のパリティグループへの割り当て例を示す概念図である。

【図6】アレイ装置における複数の記憶装置のパリティグループへの割り当て例を示す概念図である。

【図7】本発明の一実施の形態のアレイ記憶装置およびその制御方法における複数の記憶装置のパリティグループへの割り当て例を示す概念図である。

【図8】図5～図7に例示された複数の記憶装置のパリティグループへの割り当て例における、パリティグループ別接続バス比較例を示す説明図である。

【図9】本発明の一実施の形態のアレイ記憶装置およびその制御方法におけるパリティグループの初期設定方法の一例を説明するフローチャートである。

【図10】本発明の一実施の形態のアレイ記憶装置およびその制御方法におけるパリティグループの初期設定テーブルの一例を示す説明図である。

【図11】本発明の一実施の形態のアレイ記憶装置およびその制御方法におけるパリティグループを構成する記憶装置の配置の決定方法の一例を示すフローチャートである。

【図12】本発明の一実施の形態のアレイ記憶装置およびその制御方法におけるパリティグループを構成する記憶装置の配置の決定方法の一例を示すフローチャートである。

【図13】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられる接続バス競合情報テーブルの一例を示す説明図である。

【図14】本発明の一実施の形態のアレイ記憶装置およ

びその制御方法におけるパリティグループへの記憶装置の割り当て例を示す概念図である。

【図15】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられる接続バス競合情報テーブルの一例を示す説明図である。

【図16】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられる接続バス競合情報テーブルの一例を示す説明図である。

【図17】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられる接続バス競合情報テーブルの一例を示す説明図である。

【図18】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられるパリティグループ割り当てテーブルの一例を示す説明図である。

【図19】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられるチャンネル使用状況管理テーブルの一例を示す説明図である。

【図20】本発明の一実施の形態のアレイ記憶装置およびその制御方法にて用いられるチャンネル記憶装置対応テーブルの一例を示す説明図である。

【図21】本発明の一実施の形態のアレイ記憶装置およびその制御方法における障害発生時の作用の一例を示す概念図である。

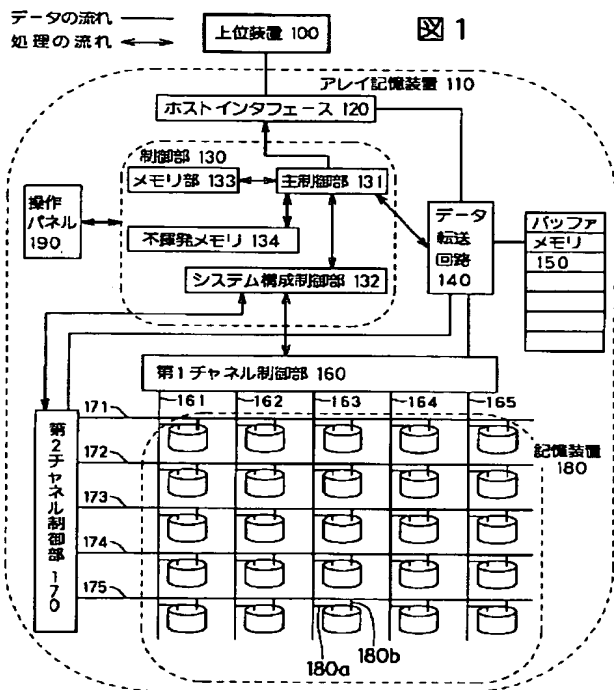
【符号の説明】

100…上位装置、110…アレイ記憶装置、120…ホストインタフェース、130…制御部、131…主制御部、132…システム構成制御部（システム構成制御手段）、133…メモリ部、134…不揮発メモリ、140…データ転送回路、150…バッファメモリ、160…第1チャンネル制御部（入出力制御部）、161～165…バス、170…第2チャンネル制御部（入出力制御部）、171～175…バス、180…記憶装置、180a…インタフェースポート、180b…インタフェースポート、190…操作パネル、200…記憶装置配置方法、201～209…記憶装置接続領域、210…記憶装置配置方法の評価、211～213…記憶装置、214…記憶装置配置方法の評価、215～217…記憶装置、218…記憶装置配置方法の評価、219～221…記憶装置、222…記憶装置配置方法の評価、223～225…記憶装置、226…記憶装置接続例、300…パリティグループ割り当て例、301…RAID定義例、302…グループ内領域割り当て例、303、304…パリティグループ、305～312…記憶装置、313…パリティグループ割り当て例、314、315…パリティグループ、316～323…記憶装置、324…パリティグループ割り当て例、325、326…パリティグループ、327～334…記憶装置、335…パリティグループ別接続バス比較例、500…パリティグループ初期設定テーブル（システム構成制御手段）、501…パリティグループ、502…RAIDレベル、

503…グループ内記憶装置数、613…接続バス競合情報テーブル（システム構成制御手段）、614…割り当て記憶装置数、617…接続バス競合情報テーブル、618…割り当て記憶装置数、619…接続バス競合情報テーブル、620…割り当て記憶装置数、621…接続バス競合情報テーブル、622…割り当て記憶装置数、700…パリティグループ割り当てテーブル、70

1…記憶装置番号、800…チャネル使用状況管理テーブル（システム構成制御手段）、900…チャネル—記憶装置対応テーブル（システム構成制御手段）、1001…バス障害時にアクセス不能な記憶装置、1002…バス障害時にアクセス可能な記憶装置、1003…バス障害時にアクセス可能な記憶装置。

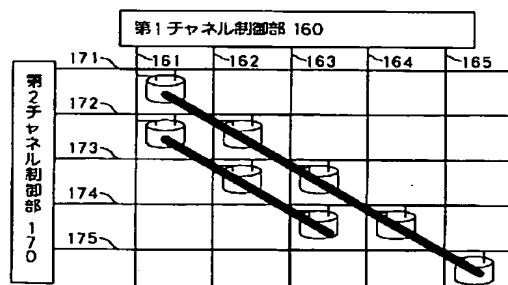
【図1】



【図4】

図4

226  
記憶装置接続例



【図2】

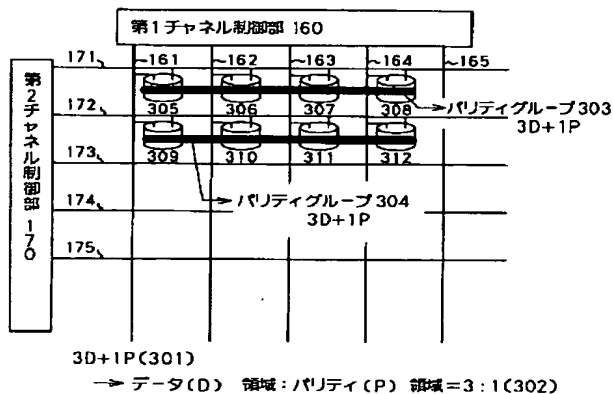
図2



【図5】

図5

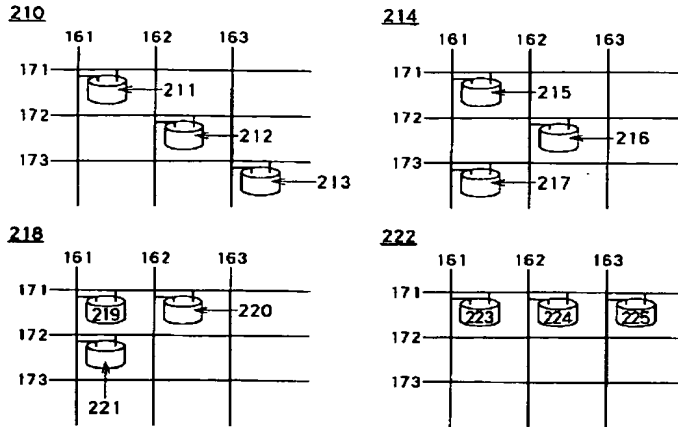
パリティグループ割り当て例 300



【図3】

図3

記憶装置配置方法の評価(210, 214, 218, 222)



【図13】

図13

613 接続バス結合情報テーブル

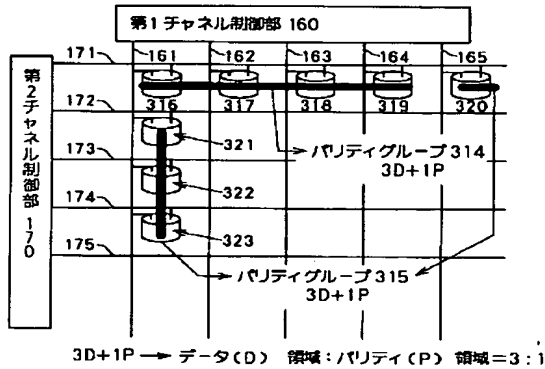
614

接続バス	割り当て記憶装置数
161	1
162	0 → 1 ← 615
:	:
172	0 → 1 ← 615
:	0

【図6】

図6

パリティグループ割り当て例 313



【図8】

図8

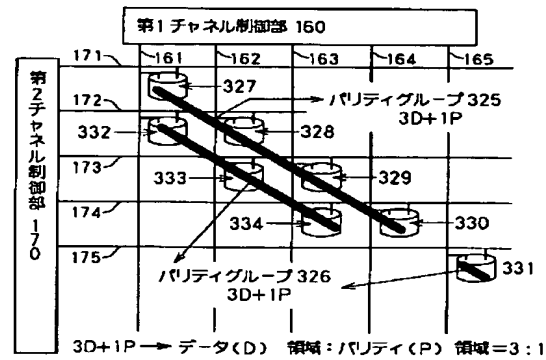
パリティグループ別接続バス比較例 335

パリティグループ番号	接続バス数	論理デバイスとして認識されているときの同時アクセス可能バス数
303	5	6
304	5	
314	5	
315	6	9
325	8	
326	8	10

【図7】

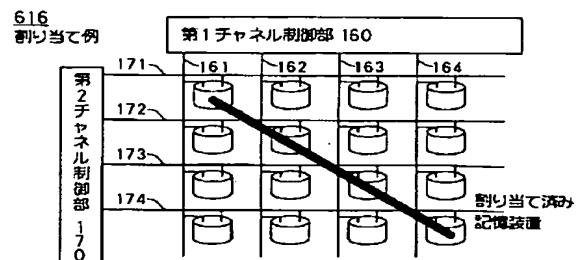
図7

パリティグループ割り当て例 324



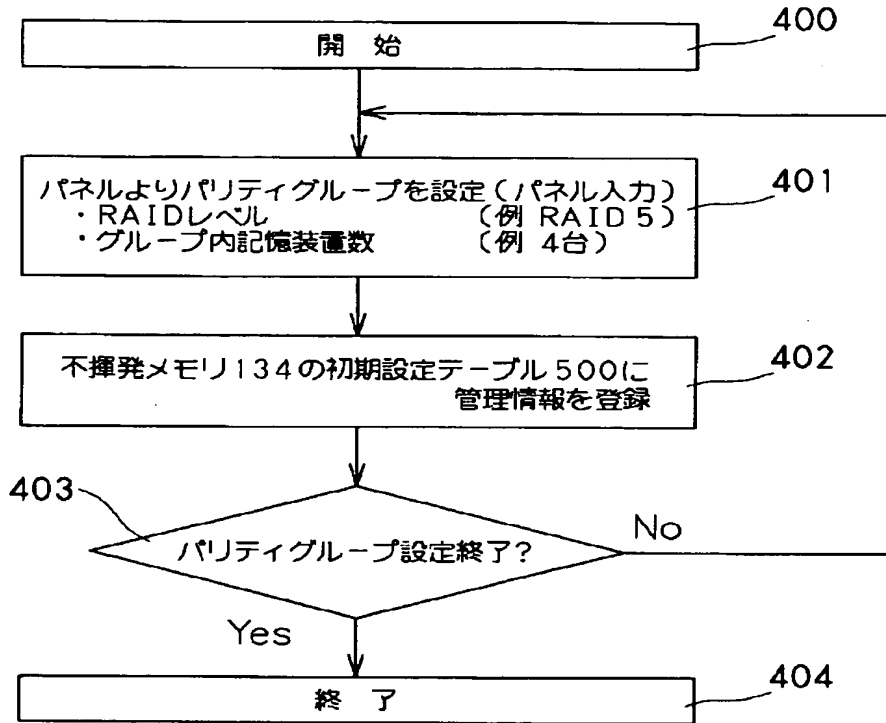
【図14】

図14



【図9】

図9



【図15】

図15

617(613) 接続/バス結合情報テーブル

接続/バス	割り当て記憶装置数
161	1
162	1
163	1
164	1
171	1
172	1
173	1
174	1

【図19】

図19

800 チャネル使用状況管理テーブル

チャネル(バス)	使用状況
161	使用中
162	未使用
163	未使用
164	使用中
:	:

【図10】

図10

500 パリティグループ初期設定テーブル(例)

項目	内容(例)
パリティグループ0	Raidレベル: RAID 5 記憶装置数: 8
パリティグループ1	Raidレベル: RAID 1 記憶装置数: 2
パリティグループ2	Raidレベル: RAID 5 記憶装置数: 3

【図16】

図16

619(613) 接続/バス結合情報テーブル

接続/バス	割り当て記憶装置数
161	2
162	1
163	1
164	1
171	1
172	2
173	1
174	1

【図17】

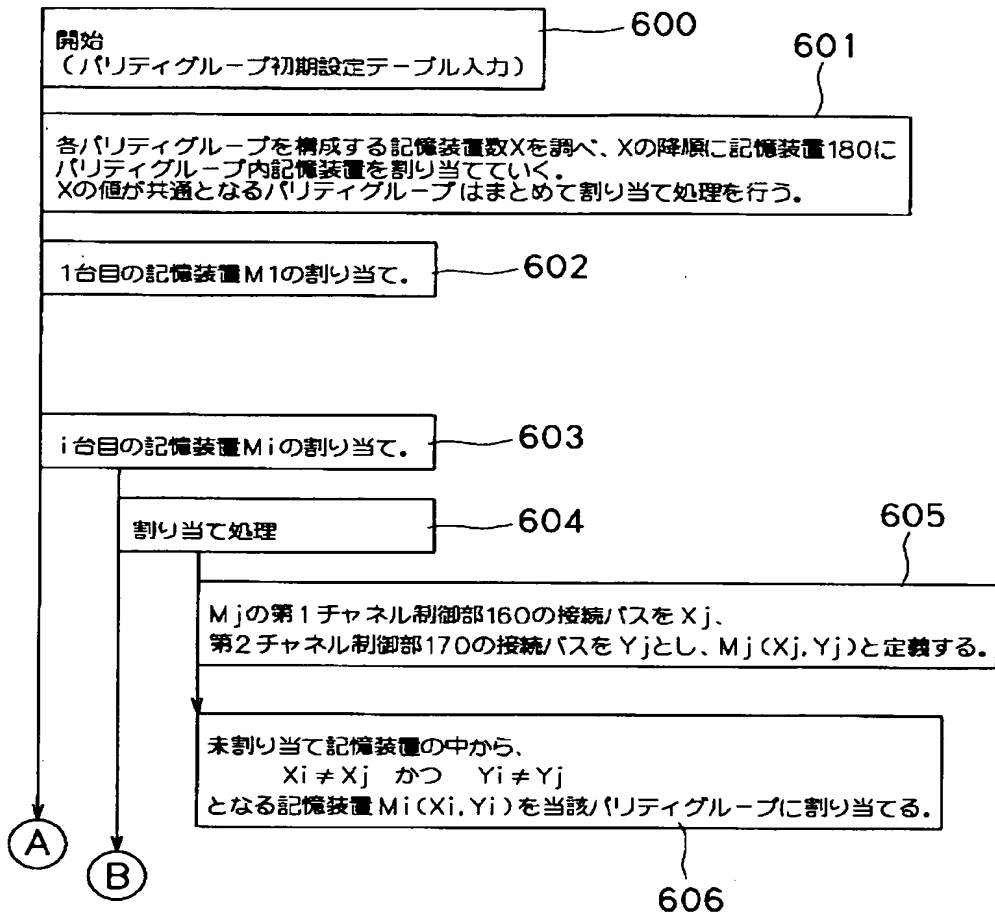
図17

621(613) 接続/バス結合情報テーブル

接続/バス	割り当て記憶装置数
161	2
162	1 → 2
163	1
164	1
171	1
172	2
173	1 → 2
174	1

【図11】

図 11



【図20】

図 20

900

チャネル-記憶装置対応テーブル

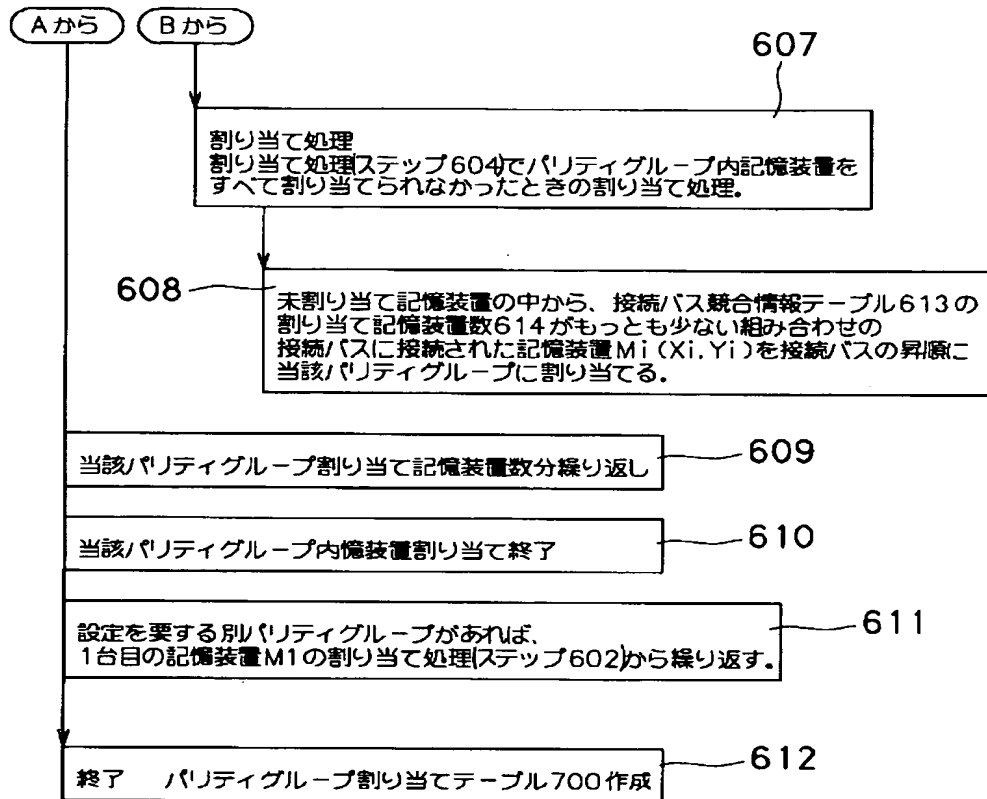
		第1チャネル制御部160					
		バス	161	162	163	164	..
制御部170	第2チャネル	171	M(161,171)	M(162,171)	M(163,171)	M(164,171)	..
		172	M(161,172)	M(162,172)	M(163,172)	M(164,172)	..
		173	M(161,173)	M(162,173)	M(163,173)	M(164,173)	..
		174	M(161,174)	M(162,174)	M(163,174)	M(164,174)	..
		:	:	:	:	:	:

記憶装置 180



【図12】

図12



【図18】

図18

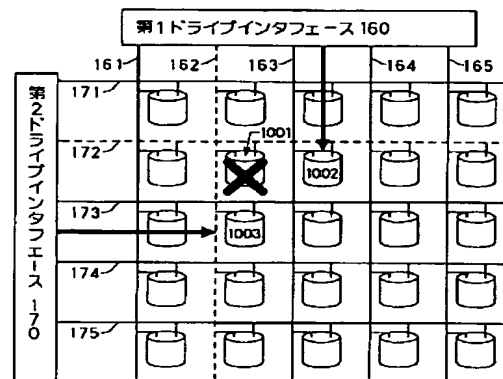
700

パリティグループ割り当てテーブル(例)

項目	名称	内容(例)
パリティグループ0	Raidレベル	RAID 5
	記憶装置数	8
	記憶装置番号	M(161,171)...
パリティグループ2	Raidレベル	RAID 5
	記憶装置数	3
	記憶装置番号	M(163,174)...
パリティグループ1	Raidレベル	RAID 1
	記憶装置数	2
	記憶装置番号	M(164,173)...

【図21】

図21



# PATENT ABSTRACTS OF JAPAN

(11)Publication number : 10-198528

(43)Date of publication of application : 31.07.1998

(51)Int.Cl.

G06F 3/06

G06F 3/06

G06F 3/06

G06F 12/16

(21)Application number : 09-001274

(71)Applicant : HITACHI LTD

(22)Date of filing : 08.01.1997

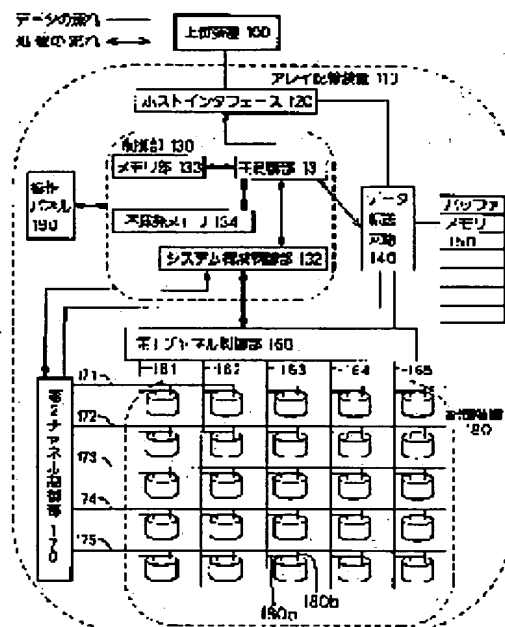
(72)Inventor : IKEDA MASANOBU  
KOBAYASHI MASAOKI  
TAKAMOTO KENICHI

## (54) ARRAY STORAGE DEVICE AND ITS CONTROLLING METHOD

## (57)Abstract:

**PROBLEM TO BE SOLVED:** To improve fault resistance performance and data input and output performance in an array storage device, and to attain the simple change of a system constitution.

**SOLUTION:** Each of plural storage devices 180 arranged like a matrix is provided with plural interface ports 180a and 180b, and connected through column directional buses 161-165 and line directional buses 171-175 with a first channel control part 160 and a second channel control part 170. Then, the plural storage devices 180 constituting the parity group of an RAID are selected so as to be exclusively connected with each of the buses 161-165 and 171-175 so that parallel access to the plural storage devices 180 in the parity group and the multiplexing of the bus connection paths can be realized. The selected result of the storage device 180 is stored in a non-volatile memory 134 of a control part 130, and the initialization/change of the parity group is operated through a control panel 190 as necessary.



## LEGAL STATUS

[Date of request for examination]

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the  
examiner's decision of rejection or application  
converted registration]

[Date of final disposal for application]

[Patent number]

[Date of registration]

[Number of appeal against examiner's decision of  
rejection][Date of requesting appeal against examiner's decision  
of rejection]

[Date of extinction of right]

